

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:50:07 ; Search time 15.665 Seconds  
(without alignments)  
370.341 Million cell updates/sec

Title: US-10-616-279-2\_COPY\_77\_91

Perfect score: 85

Sequence: 1 HSSDYSWRKNQYVS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	4	AAB82474 Human ext
2	85	100.0	290	2	AAW83329 Human min
3	85	100.0	330	4	AAE12304 Human NPG
4	85	100.0	331	2	AAW23663 Human neu
5	85	100.0	331	2	AAW70589 Adhesion-
6	85	100.0	331	2	AAW83328 Human min
7	85	100.0	331	2	AAW41721 Human PRO
8	85	100.0	331	2	AAW92460 Human NAF
9	85	100.0	331	3	AAW33465 Human PRO
10	85	100.0	331	3	AAW79561 Cancer sp
11	85	100.0	331	3	AAW44277 Human PRO
12	85	100.0	331	3	AAW95349 Human PRO
13	85	100.0	331	4	AAW93266 Human pol
14	85	100.0	331	4	AAW93324 Human pol
15	85	100.0	331	4	AAW38872 Human pol
16	85	100.0	331	4	AAW82472 Human ext
17	85	100.0	331	5	ABG61806 Prostate
18	85	100.0	331	5	AAU79944 Human Spo
19	85	100.0	331	5	ABB77393 Human spo
20	85	100.0	331	5	AAE20463 Human tum
21	85	100.0	331	6	ABO25223 Novel hum
22	85	100.0	331	6	ABU72229 Novel hum
23	85	100.0	331	6	ABU84909 Human sec
24	85	100.0	331	6	ABU61107 Human PRO
25	85	100.0	331	6	ABU80376 Human sec

## ALIGNMENTS

## RESULT 1

AAB82474

ID AAB82474 standard; peptide; 15 AA.

XX AAB82474;

XX 22-AUG-2001 (first entry)

DE Human extracellular matrix protein RGI immunogenic peptide 3C.

XX RGI; human; extracellular matrix protein; prostate cancer; metastasis;

XX tumour; benign prostatic hyperplasia; therapy; diagnosis; antitumour;

XX immunogen.

XX Homo sapiens.

XX WO200144291-A2.

XX 21-JUN-2001.

XX 15-DEC-2000; 2000WO-US033901.

XX 16-DEC-1999; 99US-0172370P.

XX 07-DEC-2000; 2000US-00732357.

XX (SCHD ) SCHERING AG.

XX Harkins R, Parkes D, Parry G, Schneider DW, Steinbrecher R;

XX WPI; 2001-398128/42.

XX Novel human extracellular matrix polypeptide, RGI, useful in research,

XX diagnosis and treatment of metastasis such as prostate cancer.

XX Claim 26; Page 48; 69pp; English.

XX The present sequence is that of immunogenic peptide 3C of human RGI (see AAB82472). RGI is a new homologue of the extracellular matrix protein family that is expressed in prostate tissue and which may be over-expressed in prostate tumours. Peptide 3C corresponds to amino acid residues 77-91 of RGI. It was selected as a potential immunogen because of its predicted position at the surface of the protein. Antisera raised against peptide 3C recognise RGI. An antibody that specifically binds to the peptide is claimed. The invention provides human RGI polypeptides, polynucleotides encoding them, and antibodies which specifically bind RGI or a polypeptide such as the present sequence. Such antibodies can be used for diagnosis and/or detection of RGI expression, or can be conjugated (e.g. as an Fv, Fab' or Fab') to a therapeutic

26	85	100.0	331	6	ABG75949	Abg75949 Human ant
27	85	100.0	331	6	ADA24775	Ada24775 Novel hum
28	85	100.0	331	6	ABO19678	AbO19678 Novel hum
29	85	100.0	331	6	ADA12436	Ada12436 Human sec
30	85	100.0	331	6	ABO19569	AbO19569 Novel hum
31	85	100.0	331	7	ADB73742	AdB73742 Human PRO
32	85	100.0	331	7	ADB76458	AdB76458 Human PRO
33	85	100.0	331	7	ADB75561	AdB75561 Prostate
34	85	100.0	331	7	ADC43884	AdC43884 Human sec
35	85	100.0	331	7	ADC61644	AdC61644 Human sec
36	85	100.0	331	7	ADC63608	AdC63608 Human sec
37	85	100.0	331	7	ADC66708	AdC66708 Human sec
38	85	100.0	331	7	ADC68832	AdC68832 Human sec
39	85	100.0	331	7	ADC62892	AdC62892 Human sec
40	85	100.0	331	7	ADC67957	AdC67957 Human sec
41	85	100.0	331	7	ADC41277	AdC41277 Human sec
42	85	100.0	331	7	ADC67332	AdC67332 Human sec
43	85	100.0	331	7	ADC62268	AdC62268 Human sec
44	85	100.0	331	7	ADC41901	AdC41901 Human sec
45	85	100.0	331	7	ADE49270	Ade49270 Human sec

CC agent, especially a cytotoxic agent, for administration to a patient for  
CC treatment of diseases characterised by RGI activity or expression, such  
CC as prostate cancer  
XX  
SQ Sequence 15 AA;  
  
Query Match 100.0%; Score 85; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4.2e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HSSDYSMWRKNQYVS 15  
| | | | | | | | | | | | | | |  
Db 1 HSSDYSMWRKNQYVS 15  
  
RESULT 2  
AAW83329  
ID AAW83329 standard; protein; 290 AA.  
AC AAW83329;  
XX  
DT 01-MAR-1999 (first entry)  
XX  
DE Human mindin-related polypeptide.  
XX  
DE Mindin; human; integrin ligand; angiogenic disease; cancer; metastasis;  
KW chronic inflammatory disorder; rheumatoid arthritis; atherosclerosis;  
KW macular degeneration, diabetic retinopathy; restenosis;  
KW Alzheimer's disease; neural disorder; tissue remodelling; therapy;  
KW diagnosis.  
XX  
OS Homo sapiens.  
XX  
PN WO9850073-A1.  
XX  
PD 12-NOV-1998.  
XX  
PF 07-MAY-1998; 98WO-US009476.  
XX  
PR 09-MAY-1997; 97US-0046106P.  
XX  
PA (SMIK ) SMITHLINE BEECHAM CORP.  
XX  
PI Jonak ZL, Trulli SH, Tsui P, Lane PA;  
XX  
DR WPI; 1999-034688/03.  
DR N-PSDB; AAV72538.  
XX  
PT A new polypeptide is the integrin ligand, human mindin - useful in  
PT therapy and diagnostic assays for diseases such as those associated with  
PT angiogenesis.  
XX  
PS Claim 14; Page 30-31; 39pp; English.  
XX  
CC This amino acid sequence was deduced from the nucleotide sequence (see  
CC AAV72538) of expressed sequence tags identified prior to the isolation of  
CC a full-length sequence (see AAV72537) encoding human mindin (see  
CC AAW83328). Claimed polypeptides have an amino acid sequence which has at  
CC least 70%, 80%, 90%, 95% or 97-77% identity to this polypeptide, or  
CC comprise the amino acid sequence of the polypeptide itself, or are  
CC encoded by a polynucleotide comprising the sequence contained in  
CC AAV72538. The invention relates to human mindin polypeptides and  
CC polynucleotides (see also AAV72537). Mindin is a novel integrin ligand  
CC suggested to have multifunctional activity in normal and disease states.  
CC Methods are provided for using mindin polypeptides and polynucleotides in  
CC the treatment and diagnosis of angiogenic diseases (cancer, cancer  
CC metastasis, chronic inflammatory disorders, rheumatoid arthritis,  
CC atherosclerosis, macular degeneration, diabetic retinopathy), restenosis,  
CC Alzheimer's disease, neural disorders and tissue remodelling. The  
CC invention also relates to methods for identifying agonists and  
CC antagonists/inhibitors and for treating conditions associated with human  
CC mindin imbalance with the identified compounds. Diagnostic assays for  
CC detecting diseases associated with inappropriate human mindin activity or

CC levels are also claimed  
XX  
SQ Sequence 290 AA;  
  
Query Match 100.0%; Score 85; DB 2; Length 290;  
Best Local Similarity 100.0%; Pred. No. 9.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HSSDYSMWRKNQYVS 15  
| | | | | | | | | | | | | | |  
Db 77 HSSDYSMWRKNQYVS 91  
  
RESULT 3  
AAE12304  
ID AAE12304 standard; protein; 330 AA.  
XX  
AC AAE12304;  
XX  
DT 03-JAN-2002 (first entry)  
XX  
DE Human NPG-1 protein.  
XX  
DE Human; NPG-1; cytostatic; gene therapy; tumour; prostate cancer; LCS;  
KW Linker Capture Subtraction; genetic alteration; nerve cell growth.  
XX  
OS Homo sapiens.  
XX  
PN US6287777-B1.  
XX  
PD 11-SEP-2001.  
XX  
PF 10-AUG-1999; 99US-00371696.  
XX  
PR 10-MAY-1996; 96US-00644326.  
PR 11-FEB-1998; 98US-00022238.  
XX  
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
XX  
PI Sytkowski AJ, Yang M;  
XX  
DR WPI; 2001-624488/72.  
DR N-PSDB; AAD20094.  
XX  
PT Detecting the presence of a nucleic acid encoding NPG-1 in a sample  
PT comprises contacting the sample with a nucleic acid probe or primer which  
PT hybridizes under stringent conditions to NPG-1 nucleic acid.  
XX  
PS Example; Fig 2; 35pp; English.  
XX  
CC The invention relates to a method termed Linker Capture Subtraction (LCS)  
CC for detecting the presence of a nucleic acid encoding NPG-1. NPG-1 gene  
CC is differentially expressed in prostate tumours. The method is used for  
CC identifying nucleic acids encoding NPG-1. NPG-1 can be used for treating  
CC a subject having a tumour, cancer, for the adhesion and outgrowth of  
CC axon, for stimulating growth of nerve cell, and regenerating nerve cells.  
CC The NPG-1 nucleic acid, protein and antibodies may be used in screening  
CC assays, detecting assays, and predictive medicine. The nucleic acids are  
CC further used to express an NPG-1 protein, to detect NPG-1 mRNA or a  
CC genetic alteration in an NPG-1 gene, and to modulate NPG-1 activity. The  
CC present sequence is human NPG-1 protein  
XX  
SQ Sequence 330 AA;  
  
Query Match 100.0%; Score 85; DB 4; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 HSSDYSMWRKNQYVS 15  
| | | | | | | | | | | | | | |  
Db 77 HSSDYSMWRKNQYVS 91

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:59:18 ; Search time 2.88177 Seconds  
(without alignments)  
500.821 Million cell updates/sec

Title: US-10-616-279-2\_COPY\_77\_91

Perfect score: 85

Sequence: 1 HSSDYSMWRKNQYVS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	54.1	438	1 ALBH	alpha-amylase (EC
2	43.5	51.2	545	2 T40207	hypothetical prote
3	42	49.4	435	2 JC7137	alpha-amylase (EC
4	42	49.4	435	2 S12625	alpha-amylase (EC
5	42	49.4	729	2 C84854	ferric-coprogen re
6	42	49.4	729	2 H90813	outer membrane rec
7	42	49.4	729	2 D85673	outer membrane rec
8	42	49.4	877	2 S58824	probable membrane
9	42	49.4	1080	2 A71485	probable pbp2-tran
10	42	49.4	3587	2 I40486	surfactin syntheta
11	41	48.2	121	2 E97103	hypothetical prote
12	41	48.2	304	2 D82189	conserved hypothet
13	41	48.2	372	2 G89921	alanine dehydrogen
14	41	48.2	410	2 S77661	conserved hypothet
15	41	48.2	432	2 G90268	conserved hypothet
16	41	48.2	805	2 T34212	hypothetical prote
17	40	47.1	135	2 C21826	alpha-amylase (EC
18	40	47.1	153	2 A21826	alpha-amylase (EC
19	40	47.1	299	2 D82880	cytosine-specific
20	40	47.1	337	2 AD3614	glycosyl transfera
21	40	47.1	423	2 T09942	alpha-amylase (EC
22	40	47.1	427	1 ALBHB	alpha-amylase (EC
23	40	47.1	429	1 J80406	sensor protein Unp
24	40	47.1	527	2 A82431	hypothetical prote
25	40	47.1	836	2 T42323	DNA-directed DNA p
26	40	47.1	1179	2 T35093	hypothetical prote
27	39	45.9	88	2 AF1023	hypothetical prote
28	39	45.9	111	2 B25159	l3k sin operon hyp
29	39	45.9	279	2 D82243	transcription regu

30 39 45.9 379 2 T50967 probable pyruvate  
31 39 45.9 425 2 S68305 gag polypeptin -  
32 39 45.9 437 2 JC7138 alpha-amylase (EC  
33 39 45.9 440 2 S14958 alpha-amylase (EC  
34 39 45.9 446 2 H90063 hypothetical prote  
35 39 45.9 494 2 T28660 probable adhesin p  
36 39 45.9 568 2 T28876 hypothetical prote  
37 39 45.9 1085 2 S55352 IFH1 protein - yea  
38 39 45.9 1150 2 T15277 hypothetical prote  
39 38 44.7 69 2 PC1257 alpha-amylase (EC  
40 38 44.7 260 2 T22990 hypothetical prote  
41 38 44.7 261 2 D86729 hypothetical prote  
42 38 44.7 296 2 T12770 probable endonucle  
43 38 44.7 347 2 T23944 hypothetical prote  
44 38 44.7 393 2 S39383 cyclin CCU1 - yea  
45 38 44.7 415 2 B84544 probable WD-40 rep

#### ALIGNMENTS

##### RESULT 1

ALBH

alpha-amylase (EC 3.2.1.1) precursor - barley

C:Species: Hordeum vulgare (barley)

C>Date: 03-Aug-1984 #sequence\_revision 03-Aug-1984 #text\_change 09-Jul-2004

C:Accession: A00846; S65602

R:Rogers, J.C.; Milliman, C.

J. Biol. Chem. 258, 8169-8174, 1983

A:Title: Isolation and sequence analysis of a barley alpha-amylase cDNA clone.

A:Reference number: A00846; MUID:83238423; PMID:6190808

A:Accession: A00846

A:Molecule type: mRNA

A:Residues: 1-438 <ROG>

A:Cross-references: UNIPROT:P00693; GB:J01236; NID:g166986; PID:AAA32929.1; PID:g166987

A:Experimental source: cv. Himalaya

R:Juge, N.; Rodenburg, K.W.; Guo, X.J.; Chaix, J.C.; Svensson, B.

FEBS Lett. 363, 299-303, 1995

A:Title: Isozyme hybrids within the protruding third loop domain of the barley alpha-amylase

A:Reference number: S65602; MUID:95255567; PMID:7737421

A:Accession: S65602

A:Molecule type: protein

A:Residues: 25-29 <JUG>

C:Comment: Production of this enzyme in barley is hormonally regulated. Germinating barl

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycogen/starch degradation

C:Superfamily: wheat alpha-amylase; alpha-amylase core homology

C:Keywords: calcium binding; glycosidase; hydrolase; monomer; polysaccharide degradation

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-438/Product: alpha-amylase #status predicted <SIG>

F:171-318/Domain: alpha-amylase core homology <AMY>

F:204,229,315/Active site: Asp, Glu, Asp #status experimental

Query Match 54.1%; Score 46; DB 1; Length 438;

Best Local Similarity 54.5%; Pred. No. 9.7;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 HSSDYSMWRKN 11

| : : : : |

DB 419 HGNDYAVWEKN 429

##### RESULT 2

T40207

hypothetical protein SPBC31F10.04c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T40207

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21913

A:Accession: T40207

A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-545 <WOO>  
A;Cross-references: UNIPROT:P87306; EMBL:Z97204; PIDN:CAB10081.1; GSPDB:GN00067; SPDB:SH  
A;Experimental source: strain 972h-; cosmid c31F10  
C;Genetics:  
A;Gene: SPDB:SPBC31F10.04c  
A;Map position: 2  
A;Introns: 30/2; 113/2

Query Match 51.2%; Score 43.5; DB 2; Length 545;  
Best Local Similarity 56.2%; Pred. No. 30;  
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 HSDYSMR-KNOYVS 15  
Db 502 HGTDSVWRCKQYSS 517

RESULT 3  
JC7137  
alpha-amylase (EC 3.2.1.1) isozyme I - rice  
N;Alternate names: 1,4-gulcan glucanohydrolase I  
C;Species: Oryza sativa (rice)  
C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 11-May-2000  
C;Accession: JC7137; PC7039  
R;Abe, R.; Yoshida, K.; Aoyagi, M.; Kasahara, S.; Ichishima, E.; Nakajima, T.  
Biosci. Biotechnol. Biochem. 63, 1329-1335, 1999  
A;Title: Characterization of chimeric enzymes constructed between two distinct alpha-amyl  
A;Reference number: JC7137; MUID:99430781; PMID:10500994  
A;Accession: JC7137  
A;Molecule type: mRNA  
A;Residues: 1-435 <AB>  
A;Accession: PC7039  
A;Molecule type: Protein  
A;Residues: 245-252 <AB2>  
C;Comment: This enzyme catalyzes the hydrolysis of internal alpha-glucosidic linkages in  
is important in germinating seeds and is present as multiple isoforms.  
C;Genetics:  
A;Gene: amy-I  
C;Superfamily: wheat alpha-amylase; alpha-amylase core homology  
C;Keywords: calcium binding; germination; glycosidase; hydrolase; seed

Query Match 49.4%; Score 42; DB 2; Length 435;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HSDYSMRK 10  
Db 415 HGKDYVWEK 424

RESULT 4  
S12625  
alpha-amylase (EC 3.2.1.1) 3D - rice  
C;Species: Oryza sativa (rice)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C;Accession: S12625; S12776; S15054; JT0945  
R;Huang, N.; Koizumi, N.; Reini, S.; Rodriguez, R.L.  
Nucleic Acids Res. 18, 7007-7014, 1990  
A;Title: Structural organization and differential expression of rice alpha-amylase genes  
A;Reference number: JT0945; MUID:91086278; PMID:2263460  
A;Accession: S12625  
A;Molecule type: DNA  
A;Residues: 1-435 <HUA>  
A;Cross-references: UNIPROT:P27933; EMBL:M59351; NID:gi169770; PIDN:AAA33895.1; PID:gi1697  
R;O'Neill, S.D.; Kumagai, M.H.; Majumdar, A.; Huang, N.; Sutliff, T.D.; Rodriguez, R.L.  
Mol. Gen. Genet. 221, 235-244, 1990  
A;Title: The alpha-amylase genes in Oryza sativa: characterization of cDNA clones and m  
A;Reference number: S12775; MUID:90318322; PMID:2370848  
A;Accession: S12776  
A;Molecule type: mRNA  
A;Residues: 1-435 <ONE>

A;Cross-references: EMBL:M24287  
R;O'Neill, S.D.; Kumagai, M.H.; Majumdar, A.; Huang, N.; Sutliff, T.D.; Rodriguez, R.L.  
submitted to the EMBL Data Library, April 1989  
A;Description: The alpha-amylase genes in Oryza sativa: Characterization of cDNA clones  
A;Reference number: S15054  
A;Accession: S15054  
A;Molecule type: mRNA  
A;Residues: 1-72, 'R', '75-136, 'R', '138-435 <ON2>  
A;Cross-references: EMBL:M24287; NID:gi169754; PIDN:AAA33886.1; PID:gi169755  
C;Genetics:  
A;Introns: 30/3; 342/3  
C;Function:  
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A;Pathway: glycogen/starch degradation  
C;Superfamily: wheat alpha-amylase; alpha-amylase core homology  
C;Keywords: glycosidase; hydrolase; polysaccharide degradation  
F;170-314/Domain: alpha-amylase core homology <AMY>  
F;203,228,311/Active site: Asp, Glu, Asp #status predicted

Query Match 49.4%; Score 42; DB 2; Length 435;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HSDYSMRK 10  
Db 415 HGKDYVWEK 424

RESULT 5  
C64854  
ferric-coprogen receptor precursor - Escherichia coli (strain K-12)  
N;Alternate names: outer membrane protein fhuE  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C;Accession: C64854; S09262; A26875; PC4409  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: C64854  
A;Accession: C64854  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-729 <BLAT>  
A;Cross-references: UNIPROT:P16869; GB:AE000210; GB:U00096; NID:gi1787332; PIDN:AAC74186  
A;Experimental source: strain K-12, substrain MGI655  
R;Sauer, M.; Hantke, K.; Braun, V.  
Mol. Microbiol. 4, 427-437, 1990  
A;Title: Sequence of the fhuE outer-membrane receptor gene of Escherichia coli K12 and p  
A;Reference number: S09262; MUID:90286919; PMID:2162465  
A;Accession: S09262  
A;Molecule type: DNA  
A;Residues: 1-362, 'C', 364-729 <SAU>  
A;Cross-references: EMBL:X17615; NID:gi14448; PIDN:CAA35616.1; PID:gi14449  
A;Experimental source: strain K-12  
R;Sauer, M.; Hantke, K.; Braun, V.  
J. Bacteriol. 169, 2044-2049, 1987  
A;Title: Ferric-coprogen receptor FhuE of Escherichia coli: processing and sequence comm  
A;Reference number: A26875; MUID:87194585; PMID:3032906  
A;Accession: A26875  
A;Molecule type: DNA  
A;Residues: 1-69 <SA2>  
R;Kobayashi, Y.; Sasamura, A.; Nishimura, A.; Kuratomi, K.  
Res. Commun. Biochem. Cell. Mol. Biol. 1, 157-170, 1997  
A;Title: A novel 14-kilodalton protein in P1,P4-bis(S'-adenosyl)tetraphosphate (AP4A)-bu  
A;Reference number: JC5685  
A;Accession: PC4409  
A;Molecule type: DNA  
A;Residues: 1-47 <KOB>  
C;Genetics:  
A;Gene: fhuB  
A;Map position: 16 min  
C;Function:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:51:02 ; Search time 14.0025 Seconds  
(without alignments)  
548.560 Million cell updates/sec

Title: US-10-616-279-2\_COPY\_77\_91

Perfect score: 85

Sequence: 1 HSSDYSWKRKNQYVS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	331	1 SPO2 HUMAN	Qbub6 homo sapien
2	82	96.5	289	2 QSKA56	Qkka56 mus musculus
3	82	96.5	330	1 SPO2 MOUSE	Q8bms2 mus musculus
4	82	96.5	330	1 SPO2 RAT	Q9wv75 rattus norv
5	82	96.5	330	2 Q8VD28	Q8vd28 mus musculus
6	63	74.1	331	2 Q42112	Q42112 brachydanio
7	55	64.7	313	2 Q6DCM4	Q6dcm4 xenopus lae
8	49	57.6	138	2 Q82UD3	Q82ud3 nitrosomona
9	49	57.6	420	2 Q22019	Q22019 cyanidiosch
10	49	57.6	446	2 Q85G84	Q85g84 cyanidiosch
11	47	55.3	390	2 Q7Y4U7	Q7y4u7 bacterioph
12	46	54.1	438	1 AMV1 HORVU	P00693 hordeum vul
13	46	54.1	438	2 Q40017	Q40017 hordeum vul
14	46	54.1	1453	2 Q75SD1	Q75sd1 ashbya goss
15	45	52.9	845	2 Q9GR11	Q9gr11 leishmania
16	44	51.8	434	2 Q81699	Q81699 avena fatua
17	44	51.8	434	2 Q87CW9	Q87cw9 xylella fas
18	44	51.8	437	2 Q81700	Q81700 avena fatua
19	44	51.8	839	2 Q8ML26	Q8ml26 drosophila
20	44	51.8	3099	2 Q7R514	Q7r514 giardia lam
21	43.5	51.2	545	1 SBB4 SCHPO	P87306 schizosacch
22	43.5	51.2	570	2 Q8D4R4	Q8d4r4 vibrio vuln
23	43.5	51.2	602	2 Q7MGB8	Q7mgb8 vibrio vuln
24	43	50.6	273	1 H1S6 MEITAC	Q8rt96 methanoearc
25	43	50.6	463	2 Q89K78	Q89kw8 bradyrhizob
26	43	50.6	809	2 Q7NAF2	Q7naf2 mycoplasma
27	42	49.4	92	2 Q6DC15	Q6dc15 brachydanio
28	42	49.4	105	2 Q73915	Q73915 bacillus ce
29	42	49.4	206	2 Q99L35	Q99l35 mus musculus
30	42	49.4	230	2 Q7T5M1	Q7t5m1 cryptophleb
31	42	49.4	239	2 Q8JR20	Q8jr20 pthorimaea

32	42	49.4	294	1 TNE7 HUMAN	O95858 homo sapien
33	42	49.4	294	2 Q6UW79	Q6uw79 homo sapien
34	42	49.4	422	2 Q6YQR6	Q6yqr6 onion yello
35	42	49.4	435	1 AM3D ORYSA	T27933 oryza sativ
36	42	49.4	436	2 Q6ZDD5	Q6zdd5 oryza sativ
37	42	49.4	442	2 Q7SDN5	Q7sdn5 neurospora
38	42	49.4	504	2 Q6L0Z8	Q6l0z8 picophilus
39	42	49.4	729	1 FHUE ECOLI	P16869 escherichia
40	42	49.4	729	2 Q7AF82	Q7af82 escherichia
41	42	49.4	729	2 Q8X8H4	Q8x8h4 escherichia
42	42	49.4	729	2 Q8CW53	Q8cw53 escherichia
43	42	49.4	734	2 Q69DL2	Q69dl2 sus scrofa
44	42	49.4	877	2 Q06593	Q06593 saccharomyc
45	42	49.4	1080	2 Q84688	Q84688 chlamydia t

#### ALIGNMENTS

#### RESULT 1

ID	SPO2 HUMAN	STANDARD;	PRT;	331 AA.
AC	Q9BUD6; Q9ULM1;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Spondin 2 precursor (MINDIN) (Differentially expressed in cancerous and noncancerous lung cells 1) (DIL-1) (UNQ435/PRO866).			
GN	Name=SPON2; Synonyms=DIL1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND VARIANTS ALA-122 AND VAL-242.			
RP	MEDLINE=39443867; PubMed=10512675; DOI=10.1006/geno.1999.5939;			
RA	Manda R., Kohno T., Matsuno Y., Takenoshita S., Kuwano H., Yokota J.;			
RT	"Identification of genes (SPON2 and C20orf2) differentially expressed between cancerous and noncancerous lung cells by mRNA differential display.";			
RT	Genomics 61:5-14(1999).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;			
RX	Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,			
RA	Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,			
RA	Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,			
RA	Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,			
RA	Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,			
RA	Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,			
RA	Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.-H., Yansura D.,			
RA	Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,			
RA	Godowski P., Gray A.;			
RT	"The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";			
RT	Genome Res. 13:2265-2270(2003).			
RL	[3]			
RN	SEQUENCE FROM N.A., AND VARIANTS ALA-122 AND VAL-242.			
RP	PubMed=14702039; DOI=10.1038/ng1285;			
RX	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,			
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hokuta T.,			
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			

RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aoteuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kunagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Takemura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RA "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs",  
RL Nat. Genet. 36:40-45(2004).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Placenta;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimod J., Schmutz J., Myers R.M.,  
RA Butterfield A.C., Krywinski W.I., Skalska U., Smallus D.E.,  
RA Scherfeld A., Schein J.E., Jones S.J.M., Marra W.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Cell adhesion protein that promote adhesion and  
CC outgrowth of hippocampal embryonic neurons. Binds directly to  
CC bacteria and their components and functions as an opsonin for  
CC macrophage phagocytosis of bacteria. Essential in the initiation  
CC of the innate immune response and represents a unique pattern-  
CC recognition molecule in the ECM for microbial pathogens (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted. Extracellular matrix (By  
CC similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in normal lung tissues but not in  
CC lung carcinoma cell lines.  
CC -!- SIMILARITY: Contains 1 spondin domain.  
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; AB027466; BAA05892.1; -;  
DR EMBL; AY358948; AAO89307.1; -;  
DR EMBL; AK074618; BAC11092.1; -;  
DR EMBL; AK074770; BAC11196.1; -;  
DR EMBL; BC002707; AAH02707.1; -;  
DR EMBL; BC036341; AAH36341.1; -;  
DR Genbank; HGNC:11253; SPON2.  
DR H-InvDB; HIX0004013; -;

DR MIM; 605918; -;  
DR InterPro; IPR009465; Spond N.  
DR InterPro; IPR000884; TSP1.  
DR Pfam; PF06468; Spond N; 1.  
DR Pfam; PF00090; TSP 1; 1.  
DR PROSITE; PS1020; SPONDIN; 1.  
DR PROSITE; PS50092; TSP1; 1.  
KW Cell adhesion; Extracellular matrix; Immune response; Polymorphism;  
KW Signal.  
FT CHAIN 1 26 Potential.  
FT DOMAIN 27 331 Spondin 2.  
FT DOMAIN 31 221 Spondin.  
FT DOMAIN 277 331 TSP type-1.  
FT VARIANT 40 40 P -> L (in dBSNP:922697).  
FT VARIANT 122 122 /FTid=VAR\_019701.  
FT VARIANT 242 242 E -> A (in dBSNP:11247975).  
FT VARIANT 242 242 L -> V (in dBSNP:2279279).  
FT VARIANT 242 242 /FTid=VAR\_019703.  
FT SEQUENCE 331 AA; 35844 MW; 418E244B893C59F4 CRC64;  
SQ SEQUENCE 331 AA; 35844 MW; 418E244B893C59F4 CRC64;  
Query Match 100.0%; Score 85; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. NO. 1.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HSSDYSMWRKNOYVS 15  
Db 77 HSSDYSMWRKNOYVS 91  
RESULT 2  
ID Q6KAS6 PRELIMINARY; PRT; 289 AA.  
AC Q6KAS6;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE MRLJ00108 protein (Fragment).  
GN Name=MRLJ00108.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,  
RA "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:  
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs  
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly  
RT Sampled from Size-Fractionated Libraries.",  
RL DNA Res. 11:167-180(2004).  
DR EMBL; AK131131; BAD21381.1; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR InterPro; IPR009465; Spond N.  
DR InterPro; IPR000884; TSP1.  
DR Pfam; PF06468; Spond N; 1.  
DR Pfam; PF00090; TSP 1; 1.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS50092; TSP1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 289 AA; 31664 MW; 7FCEB944C03021E9 CRC64;  
Query Match 96.5%; Score 82; DB 2; Length 289;  
Best Local Similarity 93.3%; Pred. NO. 3.3e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 HSSDYSMWRKNOYVS 15  
Db 35 HSSDYSMWRKNOYVS 49  
RESULT 3

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 19:00:53 ; Search time 4.10099 Seconds  
(without alignments)  
273.041 Million cell updates/sec

Title: US-10-616-279-2\_COPY\_77\_91

Perfect score: 85

Sequence: 1 HSSDYSWVRKNQYVS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	4	US-09-732-357B-10
2	85	100.0	330	3	US-09-371-696-2
3	85	100.0	331	2	US-08-799-173A-2
4	85	100.0	331	4	US-09-732-357B-2
5	85	100.0	331	4	US-09-170-042A-2
6	82	96.5	330	4	US-09-732-357B-13
7	52	61.2	299	4	US-09-311-021-202
8	44	51.8	677	4	US-09-270-767-58094
9	44	51.8	847	4	US-09-270-767-42783
10	42	49.4	25	4	US-09-270-767-58735
11	42	49.4	250	4	US-09-248-796A-15562
12	42	49.4	435	3	US-09-072-917A-9
13	42	49.4	877	4	US-09-165-396-5
14	41	48.2	149	4	US-09-270-767-47885
15	41	48.2	297	4	US-09-248-796A-25364
16	40	47.1	111	4	US-09-248-796A-21330
17	40	47.1	114	4	US-09-513-999C-5768
18	40	47.1	456	4	US-09-489-039A-8101
19	40	47.1	462	4	US-09-543-681A-5290
20	40	47.1	742	4	US-09-500-123-12
21	40	47.1	811	4	US-09-500-123-9
22	40	47.1	871	4	US-09-500-123-7
23	40	47.1	1008	4	US-09-949-016-10423
24	39	45.9	152	4	US-09-270-767-37205
25	39	45.9	152	4	US-09-270-767-52422
26	39	45.9	163	4	US-09-328-352-7800
27	39	45.9	249	4	US-09-270-767-43254

28	39	45.9	281	4	US-09-248-796A-23359	Sequence 23359, A
29	39	45.9	635	4	US-09-248-796A-16944	Sequence 16944, A
30	39	45.9	1085	1	US-08-431-080-28	Sequence 28, Appl
31	39	45.9	1085	3	US-08-938-534-28	Sequence 28, Appl
32	39	45.9	1085	3	US-09-345-294-28	Sequence 28, Appl
33	39	45.9	1169	1	US-08-542-921-2	Sequence 2, Appl
34	39	45.9	1169	2	US-08-880-685-2	Sequence 2, Appl
35	39	45.9	1169	2	US-08-880-684-2	Sequence 2, Appl
36	38.5	45.3	404	4	US-09-248-796A-19544	Sequence 19544, A
37	38	44.7	126	4	US-09-755-665-27	Sequence 27, Appl
38	38	44.7	133	4	US-09-755-665-4	Sequence 17, Appl
39	38	44.7	172	4	US-08-737-109-17	Sequence 14213, A
40	38	44.7	185	4	US-09-248-796A-14213	Sequence 6215, Ap
41	38	44.7	319	4	US-09-107-532A-6215	Sequence 9, Appl
42	38	44.7	321	3	US-08-915-795-9	Sequence 4, Appl
43	38	44.7	321	4	US-09-847-524-4	Sequence 9, Appl
44	38	44.7	321	4	US-09-296-275-9	Sequence 3, Appl
45	38	44.7	325	3	US-08-915-795-3	

#### ALIGNMENTS

RESULT 1  
US-09-732-357B-10  
; Sequence 10, Application US/09732357B  
; Patent No. 6682902  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a No. 6682902el RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSM1  
; CURRENT APPLICATION NUMBER: US/09732,357B  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-732-357B-10

Query Match 100.0%; Score 85; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSDYSWVRKNQYVS 15  
Db 1 HSSDYSWVRKNQYVS 15

RESULT 2  
US-09-371-696-2  
; Sequence 2, Application US/09371696  
; Patent No. 6287777  
; GENERAL INFORMATION:  
; APPLICANT: Sytkowski, Arthur J.  
; APPLICANT: Yang, Meiheng  
; TITLE OF INVENTION: NOVEL NPG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROSTATE  
; FILE REFERENCE: 01948/053002  
; CURRENT APPLICATION NUMBER: US/09/371,696  
; CURRENT FILING DATE: 1999-08-10  
; EARLIER APPLICATION NUMBER: US 09/022,238  
; EARLIER FILING DATE: 1998-02-11  
; EARLIER APPLICATION NUMBER: US 08/644,326  
; EARLIER FILING DATE: 1996-05-10  
; NUMBER OF SEQ ID NOS: 5

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-371-696-2

Query Match      100.0%; Score 85; DB 3; Length 330;
Best Local Similarity 100.0%; Pred. NO. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      77 HSSDYSMMRKQYVS 91

RESULT 3
US-08-799-173A-2
; Sequence 2, Application US/08799173A
; Patent No. 5871969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG,
; TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,173A
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,737
; REFERENCE/DOCKET NUMBER: PF226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-799-173A-2

Query Match      100.0%; Score 85; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. NO. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      77 HSSDYSMMRKQYVS 91

RESULT 4
US-09-732-357B-2
; Sequence 2, Application US/09732357B
; Patent No. 6682902
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
```

```
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; TITLE OF INVENTION: DNA Encoding a No. 6682902el RG-1 Polypeptide
; FILE REFERENCE: 51791AUSM1
; CURRENT APPLICATION NUMBER: US/09/732,357B
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/172,370
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-732-357B-2

Query Match      100.0%; Score 85; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. NO. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      77 HSSDYSMMRKQYVS 91

RESULT 5
US-09-170-042A-2
; Sequence 2, Application US/09170042A
; Patent No. 6759512
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg
; APPLICANT: Dillon, Patrick
; TITLE OF INVENTION: Human Neuronal Attachment Factor-1
; FILE REFERENCE: PF226D1
; CURRENT APPLICATION NUMBER: US/09/170,042A
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-170-042A-2

Query Match      100.0%; Score 85; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. NO. 4.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      77 HSSDYSMMRKQYVS 91

RESULT 6
US-09-732-357B-13
; Sequence 13, Application US/09732357B
; Patent No. 6682902
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; TITLE OF INVENTION: DNA Encoding a No. 6682902el RG-1 Polypeptide
; FILE REFERENCE: 51791AUSM1
; CURRENT APPLICATION NUMBER: US/09/732,357B
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/172,370
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 330
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 19:11:14 ; Search time 14.3719 Seconds  
(without alignments)  
436.695 Million cell updates/sec

Title: US-10-616-279-2\_COPY\_77\_91

Perfect score: 85

Sequence: 1 HSSDYSWRRKNQYVS 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	15	9	US-09-732-357A-10
2	85	100.0	15	15	US-10-616-279-10
3	85	100.0	15	16	US-10-624-884-10
4	85	100.0	15	17	US-10-895-183-10
5	85	100.0	290	15	US-10-829-952-4
6	85	100.0	330	9	US-09-903-383-2
7	85	100.0	331	9	US-09-732-357A-2
8	85	100.0	331	9	US-09-978-295A-236
9	85	100.0	331	9	US-09-938-418-8
10	85	100.0	331	9	US-09-978-697-236
11	85	100.0	331	9	US-09-978-192A-236

12	85	100.0	331	9	US-09-999-832A-236
13	85	100.0	331	10	US-09-978-189-236
14	85	100.0	331	10	US-09-978-608A-236
15	85	100.0	331	10	US-09-978-585A-236
16	85	100.0	331	10	US-09-978-191A-236
17	85	100.0	331	10	US-09-978-403A-236
18	85	100.0	331	10	US-09-978-564A-236
19	85	100.0	331	10	US-09-999-833A-236
20	85	100.0	331	10	US-09-981-915A-236
21	85	100.0	331	10	US-09-978-824-236
22	85	100.0	331	10	US-09-918-585A-236
23	85	100.0	331	10	US-09-999-834A-236
24	85	100.0	331	10	US-09-978-423A-236
25	85	100.0	331	10	US-09-978-193A-236
26	85	100.0	331	10	US-09-999-830A-236
27	85	100.0	331	10	US-09-978-757A-236
28	85	100.0	331	10	US-09-978-187B-236
29	85	100.0	331	10	US-09-978-643A-236
30	85	100.0	331	10	US-09-978-375A-236
31	85	100.0	331	10	US-09-978-298A-236
32	85	100.0	331	10	US-09-978-188A-236
33	85	100.0	331	10	US-09-978-681A-236
34	85	100.0	331	10	US-09-978-194A-236
35	85	100.0	331	10	US-09-999-829A-236
36	85	100.0	331	10	US-09-978-299A-236
37	85	100.0	331	10	US-09-978-544A-236
38	85	100.0	331	10	US-09-978-665A-236
39	85	100.0	331	10	US-09-978-802A-236
40	85	100.0	331	10	US-09-970-944-12
41	85	100.0	331	10	US-09-970-944-38
42	85	100.0	331	10	US-09-970-944-39
43	85	100.0	331	10	US-09-970-944-40
44	85	100.0	331	11	US-09-999-831A-236
45	85	100.0	331	12	US-09-978-824-236

#### ALIGNMENTS

#### RESULT 1

US-09-732-357A-10  
; Sequence 10, Application US/09732357A  
; Patent No. US20020004047A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a No. US20020004047A1 RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSM1  
; CURRENT APPLICATION NUMBER: US/09/732.357A  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-732-357A-10

Query Match 100.0%; Score 85; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSSDYSWRRKNQYVS 15

Db 1 HSSDYSWRRKNQYVS 15

#### RESULT 2

```
US-10-616-279-10
; Sequence 10, Application US/10616279
; Publication No. US20040023307A1
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; TITLE OF INVENTION: DNA Encoding a No. US20040023307A1e1 RG-1 Polypeptide
; FILE REFERENCE: 51791AUS1
; CURRENT APPLICATION NUMBER: US/10/616,279
; PRIOR FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US 60/172,370
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/732,357
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-616-279-10

Query Match      100.0%; Score 85; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      1 HSSDYSMMRKQYVS 15
|||||

RESULT 3
US-10-624-884-10
; Sequence 10, Application US/10624884
; Publication No. US20040152139A1
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; TITLE OF INVENTION: DNA Encoding a Novel RG-1 Polypeptide
; FILE REFERENCE: 51791AUS1
; CURRENT APPLICATION NUMBER: US/10/624,884
; PRIOR FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: US 60/172,370
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: US 09/732,357
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-624-884-10

Query Match      100.0%; Score 85; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      1 HSSDYSMMRKQYVS 15
|||||

US-10-616-279-2_copy_77_91.rapb
; GENERAL INFORMATION:
; APPLICANT: Harkins, Richard
; APPLICANT: Parkes, Deborah
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; TITLE OF INVENTION: RG1 Antibodies and Uses Thereof
; FILE REFERENCE: 51791BUSM1
; CURRENT APPLICATION NUMBER: US/10/895,183
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: US 60/489,032
; PRIOR FILING DATE: 2003-07-22
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-895-183-10

Query Match      100.0%; Score 85; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      1 HSSDYSMMRKQYVS 15
|||||

RESULT 5
US-10-629-952-4
; Sequence 4, Application US/10629952
; Publication No. US2004007227A1
; GENERAL INFORMATION:
; APPLICANT: Jonak, Zdenka L
; APPLICANT: Trulli, Stephen H
; APPLICANT: Tsui, Ping
; APPLICANT: Lane, Pamela A.
; TITLE OF INVENTION: INTEGRIN LIGAND, HUMAN MINDIN
; FILE REFERENCE: DDX-100US
; CURRENT APPLICATION NUMBER: US/10/629,952
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 290
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-629-952-4

Query Match      100.0%; Score 85; DB 15; Length 290;
Best Local Similarity 100.0%; Pred. No. 9.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 HSSDYSMMRKQYVS 15
Db      77 HSSDYSMMRKQYVS 91
|||||

RESULT 6
US-09-903-383-2
; Sequence 2, Application US/09903383
; Patent No. US20020137135A1
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Yang, Meiheng
; TITLE OF INVENTION: NOVEL NPG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROSTATE
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: 01948/053002
; CURRENT APPLICATION NUMBER: US/09/903,383
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/371,696
; PRIOR FILING DATE: 1999-08-10
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:50:07 ; Search time 345.675 Seconds  
(without alignments)  
370.341 Million cell updates/sec

Title: US-10-616-279-2  
Perfect score: 1760  
Sequence: 1 MENPSPAALGRKALCALLA.....NGSPCELBEEAECPDNCV 331

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1760	100.0	331	4 AAB82472	Aab82472 Human ext
2	1752	99.5	331	4 AAM93324	Aam93324 Human pol
3	1752	99.5	331	5 ABG61806	Abg61806 Prostate
4	1752	99.5	331	5 AAU79944	Aau79944 Human spo
5	1752	99.5	331	5 ABB77393	Abb77393 Human spo
6	1752	99.5	331	7 ADB75561	Adb75561 Prostate
7	1752	99.5	331	7 ADG42585	Adg42585 Human ext
8	1752	99.5	331	7 ADN38814	Adn38814 Cancer/an
9	1752	99.5	331	7 ADN39877	Adn39877 Cancer/an
10	1752	99.5	331	8 ADJ75655	Adj75655 Marker ge
11	1752	99.5	331	8 ADL30814	Adl30814 Human pro
12	1752	99.5	331	8 ADO20071	Ado20071 Human pro
13	1752	99.5	331	8 ADQ18813	Adq18813 Human sof
14	1752	99.5	422	8 ADR66362	Adr66362 Human pro
15	1752	99.5	422	8 ADR66704	Adr66704 Human pro
16	1749	99.4	331	4 AAM38872	Aam38872 Human pol
17	1749	99.4	444	4 AAM40658	Aam40658 Human pol
18	1747	99.3	331	3 AAY79561	Aay79561 Cancer sp
19	1744	99.1	331	2 AAW70589	Aaw70589 Adhesion-
20	1744	99.1	331	2 AAY41721	Aay41721 Human PRO
21	1744	99.1	331	3 AAB33465	Aab33465 Human PRO
22	1744	99.1	331	3 AAB44277	Aab44277 Human PRO
23	1744	99.1	331	3 AAY95349	Aay95349 Human PRO
24	1744	99.1	331	4 AAM93266	Aam93266 Human pol
25	1744	99.1	331	5 AAE20463	Aae20463 Human tum

26	1744	99.1	331	6 ABO25223	Abo25223 Novel hum
27	1744	99.1	331	6 ABU72229	Abu72229 Novel hum
28	1744	99.1	331	6 ABU84909	Abu84909 Human sec
29	1744	99.1	331	6 ABU61107	Abu61107 Human PRO
30	1744	99.1	331	6 ABU80376	Abu80376 Human sec
31	1744	99.1	331	6 ABG75949	Abg75949 Human ant
32	1744	99.1	331	6 ADA24775	Ada24775 Novel hum
33	1744	99.1	331	6 ABO19678	Abo19678 Novel hum
34	1744	99.1	331	6 ADA12436	Ada12436 Human sec
35	1744	99.1	331	6 ABO19569	Abo19569 Novel hum
36	1744	99.1	331	7 ADB73742	Adb73742 Human PRO
37	1744	99.1	331	7 ADB76458	Adb76458 Human PRO
38	1744	99.1	331	7 ADC43884	Adc43884 Human sec
39	1744	99.1	331	7 ADC61644	Adc61644 Human sec
40	1744	99.1	331	7 ADC83608	Adc83608 Human sec
41	1744	99.1	331	7 ADC66708	Adc66708 Human sec
42	1744	99.1	331	7 ADC68832	Adc68832 Human sec
43	1744	99.1	331	7 ADC62892	Adc62892 Human sec
44	1744	99.1	331	7 ADC67957	Adc67957 Human sec
45	1744	99.1	331	7 ADC41277	Adc41277 Human sec

## ALIGNMENTS

RESULT 1  
AAB82472  
ID AAB82472 standard; protein; 331 AA.  
XX  
AC AAB82472;  
XX  
DT 22-AUG-2001 (first entry)  
XX  
DE Human extracellular matrix protein RG1.  
XX  
KW RG1; human; extracellular matrix protein; prostate cancer; metastasis;  
KW tumour; benign prostatic hyperplasia; therapy; diagnosis; antitumour.  
XX  
OS Homo sapiens.

Key Location/Qualifiers  
FH Peptide 28..46  
FT /label= Immunogenic\_peptide  
FT /note= "specifically claimed in claim 25"  
FT Domain 31..103  
FT /label= F-spondin\_domain  
FT Peptide 77..91  
FT /label= Immunogenic\_peptide  
FT /note= "specifically claimed in claim 26"  
FT Domain 138..221  
FT /label= F-spondin\_domain  
FT Peptide 188..210  
FT /label= Immunogenic\_peptide  
FT /note= "specifically claimed in claim 27"  
FT Peptide 263..274  
FT /label= Immunogenic\_peptide  
FT /note= "specifically claimed in claim 28"  
FT Domain 278..300  
FT /label= Thrombospondin\_domain  
WO200144291-A2.

21-JUN-2001.  
15-DEC-2000; 2000WO-US033901.  
16-DEC-1999; 99US-0172370P.  
07-DEC-2000; 2000US-00732357.  
(SCHD ) SCHERING AG.  
Harkins R, Parkes D, Parry G, Schneider DW, Steinbrecher R;  
XX

DR WPI, 2001-398128/42.  
DR N-PSDB; AAF90566, AAF90567.  
XX Novel human extracellular matrix polypeptide, RGI, useful in research,  
PT diagnosis and treatment of metastasis such as prostate cancer.  
XX  
PS Claim 1; Fig 2; 69pp; English.  
XX  
CC The present sequence is that of human RGI, a new homologue of the  
CC extracellular matrix protein family, which is expressed in prostate  
CC tissue and may be over-expressed in prostate tumours. It shows homology  
CC to mindin and f-spondin proteins. This protein family is distinguished by  
CC 2 conserved spondin domains near the N-terminus and at least 1  
CC thrombospondin type 1 repeat at the C-terminus. The rgi gene (see  
CC AAF90566) was isolated from a database screening. The invention provides  
CC human RGI polypeptides (especially corresponding to amino acids 1-331, 21-  
CC 331, 27-331, 28-46 and 77-91 of RGI), polynucleotides encoding them,  
CC methods for producing the polypeptides, expression vectors, genetically  
CC engineered host cells for expression of the polypeptides, and antibodies  
CC which specifically bind RGI or its immunologically active fragment  
CC (especially a polypeptide comprising amino acid 28-46, 77-91, 188-210, or  
CC 263-274 of RGI). It also provides methods for utilising the  
CC polynucleotides, polypeptides and antibodies in research, diagnosis and  
CC therapeutic applications. Thus, the antibodies can be used for diagnosis  
CC and/or detection of RGI expression, or can be conjugated to a therapeutic  
CC agent, such as a cytotoxic agent, for administration to patient for  
CC treatment of diseases characterised by RGI activity or expression, such  
CC as prostate cancer

XX Sequence 331 AA;

Query Match 100.0%; Score 1760; DB 4; Length 331;  
Best Local Similarity 100.0%; Pred. No. 4.2e-162;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MENPSAAALGKALCALLLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
DB 1 MENPSAAALGKALCALLLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
QY 61 PLFRPPAQWSSLGAAHSSDYSMWRKNQVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
DB 61 PLFRPPAQWSSLGAAHSSDYSMWRKNQVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
QY 121 HAVFSAVAPVSGTGQTSAELEVQRHSLVSFVVRIVPSPDMFVGVDSLDLDCGDRWREQA 180  
DB 121 HAVFSAVAPVSGTGQTSAELEVQRHSLVSFVVRIVPSPDMFVGVDSLDLDCGDRWREQA 180  
QY 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTITSSPSHPANSFYYPRLKALPPIARVT 240  
DB 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTITSSPSHPANSFYYPRLKALPPIARVT 240  
QY 241 LVRLRQSPRAFIIPAPVLPSPRNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTGS 300  
DB 241 LVRLRQSPRAFIIPAPVLPSPRNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTGS 300  
QY 301 RTRVYRVQPNANGSPCPELEEEAECPDNCV 331  
DB 301 RTRVYRVQPNANGSPCPELEEEAECPDNCV 331

RESULT 2

ID AAM93324  
ID AAM93324 standard; protein; 331 AA.

XX AAM93324;

XX AAM93324;  
DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 2847.

XX Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX Homo sapiens.

OS

XX EPI130094-A2.  
XX  
XX 05-SEP-2001.  
XX  
XX 07-JUL-2000; 2000EP-00114089.  
XX  
XX 08-JUL-1999; 99JP-00194486.  
XX 11-JAN-2000; 2000JP-00118774.  
XX 02-MAY-2000; 2000JP-00183765.  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2001-524255/58.  
XX N-PSDB; AAK94244.  
XX  
XX 830 Primers useful for synthesizing full length cDNA clones and their use  
XX in genetic manipulation.  
XX  
XX Claim 8; SEQ ID NO 2847; 1380pp + Sequence Listing; English.  
XX  
XX The invention relates to primers for synthesising full length cDNA  
XX clones. 830 cDNA molecules encoding a human protein have been isolated  
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have  
XX been determined. Primers for synthesising the full length cDNA are useful  
XX for clarifying the function of the protein encoded by the cDNA. The full  
XX length clones were obtained by construction of full length enriched cDNA  
XX libraries that were synthesised by the oligo-capping method. The primers  
XX enable the production of the full length cDNA easily without any special  
XX methods. The present sequence is a polypeptide encoded by a full length  
XX human cDNA of the invention. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in CD-ROM  
XX format directly from EPO  
XX  
XX Sequence 331 AA;

Query Match 99.5%; Score 1752; DB 4; Length 331;  
Best Local Similarity 99.7%; Pred. No. 2.5e-161;  
Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MENPSAAALGKALCALLLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
DB 1 MENPSAAALGKALCALLLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
QY 61 PLFRPPAQWSSLGAAHSSDYSMWRKNQVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
DB 61 PLFRPPAQWSSLGAAHSSDYSMWRKNQVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
QY 121 HAVFSAVAPVSGTGQTSAELEVQRHSLVSFVVRIVPSPDMFVGVDSLDLDCGDRWREQA 180  
DB 121 HAVFSAVAPVSGTGQTSAELEVQRHSLVSFVVRIVPSPDMFVGVDSLDLDCGDRWREQA 180  
QY 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTITSSPSHPANSFYYPRLKALPPIARVT 240  
DB 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTITSSPSHPANSFYYPRLKALPPIARVT 240  
QY 241 LVRLRQSPRAFIIPAPVLPSPRNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTGS 300  
DB 241 LVRLRQSPRAFIIPAPVLPSPRNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTGS 300  
QY 301 RTRVYRVQPNANGSPCPELEEEAECPDNCV 331  
DB 301 RTRVYRVQPNANGSPCPELEEEAECPDNCV 331

RESULT 3

ID ABG61806  
ID ABG61806 standard; protein; 331 AA.

XX ABG61806;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:59:18 ; Search time 63.5911 Seconds  
(without alignments)  
500.821 Million cell updates/sec

Title: US-10-616-279-2  
Perfect score: 1760  
Sequence: 1 MENPSPAAALGKALCALLLA.....NGSPCELEBEAEVCVDPNCV 331

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*

- 1: pir1.\*
- 2: pir2.\*
- 3: pir3.\*
- 4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	462.5	26.3	807	2 A38152	F-spondin - rat
2	448.5	25.5	803	2 A47723	F-spondin precursor
3	401	22.8	805	2 T34212	hypothetical prote
4	122.5	7.0	741	2 I48694	probable transcrip
5	117.5	6.7	770	2 D89447	protein F57C12.1
6	112	6.4	772	2 A55004	transcription fact
7	108	6.1	123	2 S49108	TRAP-C2 protein -
8	107	6.1	808	2 T10171	phospholipase D (E
9	104.5	5.9	440	2 T24232	hypothetical prote
10	103	5.9	742	2 A49672	transcription fact
11	102	5.8	1251	2 A57293	latent transformin
12	100.5	5.7	810	2 D96566	hypothetical prote
13	100.5	5.7	2957	2 T33152	hypothetical prote
14	99.5	5.7	590	2 I46687	complement compone
15	99.5	5.7	724	2 A48569	antigen Em100 - Ei
16	99.5	5.7	812	2 T03559	phospholipase D (E
17	98	5.6	903	2 T00705	N-chimerin homolog
18	97.5	5.5	808	2 T04092	phospholipase D (E
19	97	5.5	534	2 T41081	hypothetical prote
20	97	5.5	1360	2 T33922	hypothetical prote
21	96.5	5.5	598	2 A57249	beta-galactosidase
22	96.5	5.5	712	2 A45638	immunodominant mic
23	96.5	5.5	812	2 T03402	probable phospholi
24	96.5	5.5	912	2 A54423	breivican precursor
25	96.5	5.5	1572	2 T00027	brain-specific ang
26	96	5.5	591	1 C8HUB	complement C8 beta
27	96	5.5	1584	2 T00026	brain-specific ang
28	94	5.3	809	2 T11695	phospholipase D (E
29	93	5.3	483	1 VCBPI3	minor coat protein

30	92	5.2	1306	2 S26370	MSB2 protein - Yea
31	91.5	5.2	424	2 C70651	hypothetical prote
32	91	5.2	152	2 D89753	protein Fl1c7.2 [1
33	91	5.2	862	2 T46289	hypothetical prote
34	91	5.2	937	2 D87483	ribonucleotide red
35	90	5.1	497	2 T41015	proline rich prote
36	90	5.1	692	2 AD1857	hypothetical prote
37	90	5.1	1444	2 T18856	angiogenesis inhib
38	90	5.1	1666	2 T43169	hypothetical prote
39	90	5.1	2265	1 FNBO	fibronectin - bovi
40	90	5.1	3027	2 JQ1917	polyprotein - pars
41	89.5	5.1	919	2 T32541	unc-5 protein - Ca
42	89.5	5.1	947	1 B44294	unc-5 protein, lon
43	89	5.1	649	2 D96025	probable adenylate
44	89	5.1	697	2 T03834	nuclear disributi
45	89	5.1	1257	2 T09493	period protein hom

ALIGNMENTS

RESULT 1

A38152

F-spondin - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A38152

R:Klar, A.; Baldassare, M.; Jessell, T.M.

Cell 69, 95-110, 1992

A:Title: F-spondin: a gene expressed at high levels in the floor plate encodes a secretor

A:Reference number: A38152; MUID:92208952; PMID:1555244

A:Accession: A38152

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-807 <KLA>

A:Cross-references: UNIPROT:P35446; GB:M8469; NID:G204176; PIDN:AAA41174.1; PID:G204177

A:Experimental source: embryo floor plate

C:Note: sequence extracted from NCBI backbone (NCBIN:90877, NCBIP:90878)

C:Superfamily: F-spondin; thrombospondin type 1 repeat homology

F;441-495/Domain: thrombospondin type 1 repeat homology <THR2>

F;500-555/Domain: thrombospondin type 1 repeat homology <THR3>

F;557-611/Domain: thrombospondin type 1 repeat homology <THR4>

F;613-666/Domain: thrombospondin type 1 repeat homology <THR1>

F;667-721/Domain: thrombospondin type 1 repeat homology <THR5>

F;753-807/Domain: thrombospondin type 1 repeat homology <THR6>

Query Match 26.3%; Score 462.5; DB 2; Length 807;

Best Local Similarity 34.2%; Pred. No. 1.6e-28;

Matches 116; Conservative 49; Mismatches 137; Indels 37; Gaps 12;

QY 9 ALGKALCALLLTLGAGQPLGGESICSAQAPAKYSITFTGKWSQTAFFKQYPLFRPPAQ 68

Db 176 SLTKKLCEQDPTLDGVTDTRPI---LDCCACGTAKYRLTFYGNWSEKTHPKDYP--RRANH 230

QY 69 WSLLLGAAHSSDYSMWRKQYVNGLRDFAERGEAWALMKEI-EAAGEALQSVHAFSAP 127

Db 231 WSAIIGSHSKNYLWEYGGYASBQVVAELGSPVVMEEIRQQQDEVLTVTKAKAQP 290

QY 128 -----AVPSGTGTSAELEVQRHSLVSFVVRIVPSPDWFVGVDSLDLDCGD-RWRE 178

Db 291 SWQPVNVRAP-----SAEFSVDRTHLSMFLTMWGSPDNNVGLSAEDLCTKECGWVQ 344

QY 179 QAALDLYPYDAGTDSGFTSSNFATIPQDVTVEITSSPSHPANFYPRLLKALPIAR 238

Db 345 KVVQDLIPWDAGTDSGVTVESPNKPTIPQEKIRPLTSL--DHPQSPFYDEGGSSITQVAR 402

QY 239 VTLVRL-RQSPRAFIPAPVLPSPRDNEIVDSASVPE-----TPLDCEVLSWSWGLCGG 291

Db 403 VWIERIARGEQCNIYPDNV---DDIVADLA--PEEKEDDDTPTETCIYSNWSFWSACSS 456

QY 292 HCGRLGTGKSRTRVVRVQPPANNNGSPCELEBEAEVCVDPNC 330

Db 457 STCEKGRMRQRMKQAQ-LDLSVPCFDTQDFQPCMGPGC 494



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 18:51:02 ; Search time 308.988 Seconds  
(without alignments)  
548.560 Million cell updates/sec

Title: US-10-616-279-2

Perfect score: 1760

Sequence: 1 MENPSPAAALGKALCALLA.....NGSPCELEBAECVDPNCV 331

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	%	Query	Length	DB	ID	Description
1	1744	99.1	331	1	SPO2_HUMAN			Q9bud6 homo sapien
2	1506.5	85.6	330	1	SPO2_RAT			Q9w775 rattus norv
3	1494.5	84.9	330	1	SPO2_MOUSE			Q8bms2 mus musculu
4	1488.5	84.6	330	2	Q8VD28			Q8vd28 mus musculu
5	1403	79.7	289	2	Q6KAS6			Q6kas6 mus musculu
6	1113.5	63.3	313	2	Q6DCM4			Q6dcm4 xenopus lae
7	1108	63.0	331	2	Q42112			Q42112 brachydanio
8	871	49.5	334	2	Q42111			Q42111 brachydanio
9	522.5	29.7	601	2	Q9V746			Q9v746 drosophila
10	512	29.1	598	2	O02029			O02029 drosophila
11	470.5	26.7	808	2	Q42113			Q42113 brachydanio
12	466	26.5	729	2	Q69Z27			Q69z27 mus musculu
13	464.5	26.4	802	1	SPO1_CHICK			Q9w770 gallus gall
14	462.5	26.3	807	1	SPO1_MOUSE			Q8vcc9 mus musculu
15	462.5	26.3	807	1	SPO1_RAT			P35446 rattus norv
16	460.5	26.2	807	1	SPO1_BOVIN			Q9glx9 bos taurus
17	460.5	26.2	807	1	SPO1_HUMAN			Q9hcb6 homo sapien
18	457	26.0	898	2	Q76822			Q76822 branchiost
19	448.5	25.5	803	1	SPO1_XENLA			P35447 xenopus lae
20	447	25.4	628	2	Q7KRFP			Q7krf4 drosophila
21	447	25.4	763	2	Q9XZD0			Q9xzd0 drosophila
22	443	25.2	803	2	Q42114			Q42114 brachydanio
23	441.5	25.1	873	2	Q7KR42			Q7kr42 drosophila
24	440.5	25.0	951	2	Q7Q082			Q7q082 anopheles g
25	436	24.8	608	2	Q7PZ75			Q7pz75 anopheles g
26	424.5	24.1	839	2	Q8ML26			Q8ml26 drosophila
27	420.5	23.9	216	2	Q9H711			Q9h711 homo sapien
28	401	22.8	819	2	Q19305			Q19305 caenorhabdi
29	306	17.4	861	2	Q8ML27			Q8ml27 drosophila
30	280.5	15.9	549	2	Q8T988			Q8t988 drosophila
31	271.5	15.4	461	2	Q95822			Q95822 drosophila

32	225.5	12.8	92	2	Q6DC15	Q6dc15 brachydanio
33	149.5	8.5	1107	2	Q8BHP3	Q8bhp3 mus musculu
34	149	8.5	1607	2	Q6F4U0	Q6f4u0 mus musculu
35	144.5	8.2	1536	2	Q9C0I4	Q9c0i4 homo sapien
36	140.5	8.0	238	2	Q69HT6	Q69ht6 ciona intes
37	134	7.6	1502	2	Q9UPZ6	Q9upz6 homo sapien
38	134	7.6	1668	2	Q6ZU6	Q6zu6 mus musculu
39	131	7.4	856	2	Q86PQ8	Q86pq8 cryptospori
40	130	7.4	3869	2	Q86PQ3	Q86pq3 cryptospori
41	128	7.3	687	2	Q23729	Q23729 cryptospori
42	127	7.2	921	2	Q969A3	Q969a3 branchiost
43	125	7.1	660	2	Q23832	Q23832 cryptospori
44	123	7.0	243	2	Q8BFU0	Q8bfu0 m mus muscu
45	122.5	7.0	437	2	Q7YF59	Q7yfy59 cryptospori

#### ALIGNMENTS

#### RESULT 1

ID	SPO2_HUMAN	STANDARD;	PRT;	331 AA.
AC	Q9BUD6; Q9ULW1;			
DT	25-OCT-2004 (Rel. 45, Created)			
DT	25-OCT-2004 (Rel. 45, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Spondin 2 precursor (Mindin) (Differentially expressed in cancerous and noncancerous lung cells 1) (DIL-1) (UNQ435/PRO866).			
GN	Name=SPON2; Synonyms=DIL1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND VARIANTS ALA-122 AND VAL-242.			
RX	MEDLINE=99443867; PubMed=10512675; DOI=10.1006/geno.1999.5939;			
RT	Manda R., Kohno T., Matsuno Y., Takenoshita S., Kuwano H., Yokota J.;			
RT	"Identification of genes (SPON2 and C20orf2) differentially expressed			
RT	between cancerous and noncancerous lung cells by mRNA differential			
RT	display.";			
RL	Genomics 61:5-14 (1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;			
RA	Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,			
RA	Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,			
RA	Eaton D., Foster J., Grimaldi C., Gu O., Hass P.E., Heldens S.,			
RA	Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,			
RA	Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,			
RA	Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,			
RA	Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.-H., Yansura D.,			
RA	Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,			
RA	Godowski P., Gray A.;			
RT	"The secreted protein discovery initiative (SPDI), a large-scale			
RT	effort to identify novel human secreted and transmembrane proteins: a			
RT	bioinformatics assessment.";			
RL	Genome Res. 13:2265-2270 (2003).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS ALA-122 AND VAL-242.			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Oiyashiki M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			
RA	Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,			
RA	Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,			
RA	Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,			
RA	Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,			
RA	Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,			
RA	Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,			
RA	Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,			
RA	Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,			

RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
RA Imose N., Mueashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
RA Moriya S., Momiya M., Satoh N., Takami S., Terashima Y., Suzuki O.,  
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
RA Fujimori Y., Koniya M., Tashiro H., Tanigami A., Fujiwara T.,  
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
RA Nakai K., Yada I., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
RT "Complete sequencing and characterization of 21,243 full-length human  
RT cDNAs.";  
RL Nat. Genet. 36:40-45(2004).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Placenta;  
RX MEDLINE=22388257; PubMed12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Prange C.,  
RA Raha S.S., McQuillan N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., Moqlwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences";  
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Cell adhesion protein that promotes adhesion and  
CC outgrowth of hippocampal embryonic neurons. Binds directly to  
CC bacteria and their components and functions as an opsonin for  
CC macrophage phagocytosis of bacteria. Essential in the initiation  
CC of the innate immune response and represents a unique pattern-  
CC recognition molecule in the ECM for microbial pathogens (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted. Extracellular matrix (By  
CC similarity).  
CC -!- TISSUE SPECIFICITY: Expressed in normal lung tissues but not in  
CC lung carcinoma cell lines.  
CC -!- SIMILARITY: Contains 1 spondin domain.  
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AB027466; BAA85892.1; -;  
CC EMBL: AY358948; AAQ89307.1; -;  
CC EMBL: AK074618; BAC11092.1; -;  
CC EMBL: AK074770; BAC11196.1; -;  
CC EMBL: BC027007; AAH02707.1; -;  
CC EMBL: BC036341; AAH36341.1; -;  
CC EMBL: HGNC:11253; SPON2.  
CC H-InvDB: HIX0004013; -;

DR MIM: 605918; -;  
DR InterPro: IPR009465; Spond N.  
DR InterPro: IPR000884; TSP1.  
DR Pfam: PF06468; Spond N; 1.  
DR Pfam: PF00090; TSP 1; 1.  
DR PROSITE: PS1020; SPONDIN; 1.  
DR PROSITE: PS50092; TSP1; 1.  
KW Cell adhesion; Extracellular matrix; Immune response; Polymorphism;  
KW Signal.  
FT SIGNAL. 1 26 Potential.  
FT CHAIN 27 331 Spondin 2.  
FT DOMAIN 21 221 Spondin.  
FT VARIANT 277 331 TSP type-1.  
FT VARIANT 40 40 P -> L (in dbSNP:922697).  
FT VARIANT 122 122 E -> A (in dbSNP:11247975).  
FT VARIANT 242 242 L -> V (in dbSNP:2279279).  
FT VARIANT 242 242 L -> V (in dbSNP:019702).  
FT VARIANT 242 242 L -> V (in dbSNP:019703).  
FT SEQUENCE 331 AA; 35844 MW; 418E244B893C59F4 CRC64;  
Query Match 99.1%; Score 1744; DB 1; Length 331;  
Best Local Similarity 99.1%; Pred. No. 2.5e-128;  
Matches 328; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MENPSPAALGKALCALLLTLGAAGQPLGSGICSGAGAPAKYSITFTGKWSQTAFPKQY 60  
DB 1 MENPSPAALGKALCALLLTLGAAGQPLGSGICSGARAPAKYSITFTGKWSQTAFPKQY 60  
QY 61 PLFRPPAQSLLGAHSSDYSMWRKQYVNSGLRDPFAERGEAWALMKEIEAAGALQSV 120  
DB 61 PLFRPPAQSLLGAHSSDYSMWRKQYVNSGLRDPFAERGEAWALMKEIEAAGALQSV 120  
QY 121 HAVFAPAVPGTGTGTAELVQRHSLVSFVVRIVPSDFVGVDSLDLDCDGRWRQEA 180  
DB 121 HEVFSAPAVPGTGTGTAELVQRHSLVSFVVRIVPSDFVGVDSLDLDCDGRWRQEA 180  
QY 181 ALDLYPYDAGTDSGTFTSSPNFATIPQDVTTEITSSPSHPANSFYPRKALPFIARVT 240  
DB 181 ALDLYPYDAGTDSGTFTSSPNFATIPQDVTTEITSSPSHPANSFYPRKALPFIARVT 240  
QY 241 LVRLQSPRAFTIPAPVLPFRDNEIVDSASVPTPLDCEVLSWSSWGLCGHCGRLGTS 300  
DB 241 LVRLQSPRAFTIPAPVLPFRDNEIVDSASVPTPLDCEVLSWSSWGLCGHCGRLGTS 300  
QY 301 RTRYVRVOPANNQSPCELBEEAECPDNCV 331  
DB 301 RTRYVRVOPANNQSPCELBEEAECPDNCV 331  
RESULT 2  
SPO2 RAT  
ID SPO2 RAT STANDARD; PRT; 330 AA.  
AC QSW75;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Spondin 2 precursor (Mandin).  
GN Names=Spond2;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=99339921; PubMed=10409509;  
RA Feinstein Y., Borrell V., Garcia C., Burstin-Cohen T., Tzarfaty V.,  
RA Frumkin A., Nose A., Okamoto H., Higashijima S., Soriano A., Klar A.;  
RT "P-spondin and mindin: two structurally and functionally related genes  
RT expressed in the hippocampus that promote outgrowth of embryonic  
RT hippocampal neurons";  
RL Development 126:3637-3648 (1999).



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2005, 19:00:53 ; Search time 90.4951 Seconds  
(without alignments)  
273.041 Million cell updates/sec

Title: US-10-616-279-2

Perfect score: 1760

Sequence: 1 MENPSPAALGKALCALLA.....NGSPCELEEEACVDPNCV 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/protdata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/protdata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/protdata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/protdata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/protdata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/protdata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	331	4	US-09-732-357B-2
2	1742	99.0	331	2	US-08-799-173A-2
3	1742	99.0	331	4	US-09-170-042A-2
4	1551.5	88.2	330	3	US-09-371-696-2
5	1506.5	85.6	330	4	US-09-732-357B-13
6	1101.5	62.6	299	4	US-09-311-021-202
7	464.5	26.4	802	1	US-07-862-021B-12
8	464.5	26.4	802	1	US-08-313-288B-12
9	464.5	26.4	802	5	PCT-US93-03164-12
10	462.5	26.3	392	2	US-08-799-173A-7
11	462.5	26.3	392	4	US-09-170-042A-7
12	462.5	26.3	807	1	US-07-862-021B-10
13	462.5	26.3	807	1	US-08-313-288B-10
14	462.5	26.3	807	4	US-09-132-769-5
15	462.5	26.3	807	5	PCT-US93-03164-10
16	460.5	26.2	787	4	US-09-825-294-207
17	460.5	26.2	787	4	US-09-970-966-207
18	460.5	26.2	807	4	US-09-132-769-1
19	460.5	26.2	807	4	US-09-132-769-3
20	460.5	26.2	807	4	US-09-640-173-186
21	460.5	26.2	807	4	US-09-713-550-186
22	460.5	26.2	807	4	US-09-825-294-186
23	460.5	26.2	807	4	US-09-970-966-186
24	438.5	24.9	819	4	US-09-270-767-42963
25	422.5	24.0	677	4	US-09-270-767-58094
26	422.5	24.0	847	4	US-09-270-767-42783
27	420.5	23.9	132	3	US-09-022-238-2

28	331	18.8	568	1	US-07-862-021B-14	Sequence 14, Appl
29	331	18.8	588	5	PCT-US93-03164-14	Sequence 14, Appl
30	309	17.6	53	2	US-08-799-173A-18	Sequence 18, Appl
31	309	17.6	53	4	US-09-170-042A-19	Sequence 19, Appl
32	189	10.7	37	3	US-09-022-238-3	Sequence 3, Appli
33	181	10.3	37	3	US-09-371-696-3	Sequence 3, Appli
34	130.5	7.4	56	1	US-07-862-021B-19	Sequence 19, Appl
35	130.5	7.4	56	5	PCT-US93-03164-19	Sequence 19, Appl
36	128.5	7.3	50	2	US-08-799-173A-14	Sequence 14, Appl
37	128.5	7.3	50	4	US-09-170-042A-14	Sequence 14, Appl
38	121.5	6.9	52	2	US-08-799-173A-12	Sequence 12, Appl
39	121.5	6.9	52	4	US-09-170-042A-12	Sequence 12, Appl
40	116	6.6	23	4	US-09-732-357B-11	Sequence 11, Appl
41	110.5	6.3	149	4	US-09-270-767-32202	Sequence 32202, A
42	110.5	6.3	149	4	US-09-270-767-47419	Sequence 47419, A
43	108	6.1	19	4	US-09-732-357B-9	Sequence 9, Appli
44	103	5.9	229	4	US-09-894-912A-25	Sequence 25, Appl
45	101	5.7	160	4	US-09-894-912A-14	Sequence 14, Appl

#### ALIGNMENTS

RESULT 1  
US-09-732-357B-2  
; Sequence 2, Application US/09732357B  
; Patent No. 6682902  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a No. 6682902el RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSM1  
; CURRENT APPLICATION NUMBER: US/09/732,357B  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-732-357B-2

Query Match	100.0%	Score 1760;	DB 4;	Length 331;
Best Local Similarity	100.0%	Pred. No. 8.2e-166;	Mismatches 0;	Indels 0;
Matches 331;	Conservative 0;			Gaps 0;
QY	1	MENPSPAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAFPKQY	60	
DB	1	MENPSPAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAFPKQY	60	
QY	61	PLFRPPAQWSSLLGAHSSDYSMWRKQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV	120	
DB	61	PLFRPPAQWSSLLGAHSSDYSMWRKQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV	120	
QY	121	HAVFSAPVPSGGTGSAELEVQRHSLVSFVVRIVPSDFVGVDSLDCDGRWREQA	180	
DB	121	HAVFSAPVPSGGTGSAELEVQRHSLVSFVVRIVPSDFVGVDSLDCDGRWREQA	180	
QY	181	ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHSPANSFYYPRLKALPPIARVT	240	
DB	181	ALDLYPYDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHSPANSFYYPRLKALPPIARVT	240	
QY	241	LVLRLQSPRAFIAPPVLPVPSRDNIEVDASVPETPLDCEVSLWSSWGLCGHCGRLGTSK	300	
DB	241	LVLRLQSPRAFIAPPVLPVPSRDNIEVDASVPETPLDCEVSLWSSWGLCGHCGRLGTSK	300	
QY	301	RTEYVSVQPNNGSPCELEEEACVDPNCV	331	

```
Db      301 RTRVVRVQPNNGSPCPPELEEEAEACVPDNCV 331

RESULT 2
US-08-799-173A-2
; Sequence 2, Application US/08799173A
; Patent No. 5871969
; GENERAL INFORMATION:
; APPLICANT: HASTINGS, GREGG,
; APPLICANT: PATRICK J. DILLON
; TITLE OF INVENTION: HUMAN NEURONAL ATTACHMENT FACTOR-1
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 11-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROOKES, ANDERS A.
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-799-173A-2

Query Match      99.0%; Score 1742; DB 2; Length 331;
Best Local Similarity 99.4%; Pred. No. 5e-164;
Matches 329; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 MENPSAAALGKALCALLATLGAAGQPLGGESICSAAPAKYSITFTGKWSQTAPFKQY 60
      |||
Db      1 MENPSAAALGKALCALLATLGAAGQPLGGESICSAARALAKYSITFTGKWSQTAPFKQY 60
      |||

QY      61 PLFRPPAQMSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120
      |||
Db      61 PLFRPPAQMSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120
      |||

QY      121 HAVFSAPAVPSGTGQTSAELEVQRHSLVSFVVRIVPSDPDFVGVDSLDLDCGDRWREQA 180
      |||
Db      121 HAVFSAPAVPSGTGQTSAELEVQRHSLVSFVVRIVPSDPDFVGVDSLDLDCGDRWREQA 180
      |||

QY      181 ALDLYPYDAGTDSGFTSSPNFATIPQDVTTEITSSSPSHPANSFYPRKALPPIARVT 240
      |||
Db      181 ALDLYPYDAGTDSGFTSSPNFATIPQDVTTEITSSSPSHPANSFYPRKALPPIARVT 240
      |||

QY      241 LVLRQSPRAFIPAPVLPSPRDNIEIVDSASVPETPLDCEVSLWSSWGLCGHCGRLGTSK 300
      |||
Db      241 LVLRQSPRAFIPAPVLPSPRDNIEIVDSASVPETPLDCEVSLWSSWGLCGHCGRLGTSK 300
      |||

QY      301 RTRVVRVQPNNGSPCPPELEEEAEACVPDNCV 331
      |||
Db      301 RTRVVRVQPNNGSPCPPELEEEAEACVPDNCV 331
      |||

RESULT 4
US-09-371-696-2
; Sequence 2, Application US/09371696
; Patent No. 6287777
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Yang, Meiheng
; TITLE OF INVENTION: NOVEL NPG-1 GENE THAT IS DIFFERENTIALLY EXPRESSED IN PROSTATE
; FILE OF INVENTION: TUMORS
; FILE REFERENCE: 01948/053002
; CURRENT APPLICATION NUMBER: US/09/371,696
; EARLIER FILING DATE: 1999-08-10
; EARLIER APPLICATION NUMBER: US 09/022,238
; EARLIER FILING DATE: 1998-02-11
; EARLIER APPLICATION NUMBER: US 08/644,326
; EARLIER FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-371-696-2

Query Match      88.2%; Score 1551.5; DB 3; Length 330;
Best Local Similarity 87.6%; Pred. No. 3.6e-145;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: November 14, 2005, 19:11:14 ; Search time 317.14 Seconds  
(without alignments)  
436.695 Million cell updates/sec

Title: US-10-616-279-2  
Perfect score: 1760  
Sequence: 1 MENPSPAALGKALCALLLA.....NGSPCELEBEAECPDNCV 331

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10F\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1760	100.0	331	9	US-09-732-357A-2
2	1760	100.0	331	15	US-10-616-279-2
3	1760	100.0	331	16	US-10-624-884-2
4	1760	100.0	331	17	US-10-895-183-2
5	1752	99.5	331	10	US-09-970-944-12
6	1752	99.5	331	14	US-10-205-823-385
7	1752	99.5	331	15	US-10-295-027-132
8	1752	99.5	331	15	US-10-295-027-1195
9	1752	99.5	331	16	US-10-723-860-1632
10	1752	99.5	331	18	US-10-631-467-907
11	1752	99.5	331	18	US-10-631-467-907

12	1752	99.5	331	20	US-11-051-454-385	Sequence 385, App
13	1747	99.3	331	18	US-10-929-973-2	Sequence 2, Appli
14	1747	99.3	349	18	US-10-919-215-1	Sequence 1, Appli
15	1744	99.1	331	9	US-09-978-295A-236	Sequence 236, App
16	1744	99.1	331	9	US-09-938-418-8	Sequence 8, Appli
17	1744	99.1	331	9	US-09-978-697-236	Sequence 236, App
18	1744	99.1	331	9	US-09-978-192A-236	Sequence 236, App
19	1744	99.1	331	9	US-09-999-832A-236	Sequence 236, App
20	1744	99.1	331	10	US-09-978-189-236	Sequence 236, App
21	1744	99.1	331	10	US-09-978-608A-236	Sequence 236, App
22	1744	99.1	331	10	US-09-978-585A-236	Sequence 236, App
23	1744	99.1	331	10	US-09-978-191A-236	Sequence 236, App
24	1744	99.1	331	10	US-09-978-403A-236	Sequence 236, App
25	1744	99.1	331	10	US-09-978-564A-236	Sequence 236, App
26	1744	99.1	331	10	US-09-999-833A-236	Sequence 236, App
27	1744	99.1	331	10	US-09-981-915A-236	Sequence 236, App
28	1744	99.1	331	10	US-09-978-824-236	Sequence 236, App
29	1744	99.1	331	10	US-09-918-585A-236	Sequence 236, App
30	1744	99.1	331	10	US-09-999-834A-236	Sequence 236, App
31	1744	99.1	331	10	US-09-978-423A-236	Sequence 236, App
32	1744	99.1	331	10	US-09-978-193A-236	Sequence 236, App
33	1744	99.1	331	10	US-09-999-830A-236	Sequence 236, App
34	1744	99.1	331	10	US-09-978-757A-236	Sequence 236, App
35	1744	99.1	331	10	US-09-978-187B-236	Sequence 236, App
36	1744	99.1	331	10	US-09-978-643A-236	Sequence 236, App
37	1744	99.1	331	10	US-09-978-375A-236	Sequence 236, App
38	1744	99.1	331	10	US-09-978-298A-236	Sequence 236, App
39	1744	99.1	331	10	US-09-978-188A-236	Sequence 236, App
40	1744	99.1	331	10	US-09-978-681A-236	Sequence 236, App
41	1744	99.1	331	10	US-09-978-194A-236	Sequence 236, App
42	1744	99.1	331	10	US-09-999-829A-236	Sequence 236, App
43	1744	99.1	331	10	US-09-978-299A-236	Sequence 236, App
44	1744	99.1	331	10	US-09-978-544A-236	Sequence 236, App
45	1744	99.1	331	10	US-09-978-665A-236	Sequence 236, App

## ALIGNMENTS

## RESULT 1

US-09-732-357A-2  
; Sequence 2, Application US/09732357A  
; Patent No. US20020004047A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a No. US20020004047A1el RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSM1  
; CURRENT APPLICATION NUMBER: US/09/732,357A  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 331  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-732-357A-2

Query Match 100.0%; Score 1760; DB 9; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1.6e-150;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MENPSPAALGKALCALLLATLGAAGQPLGSGESCSAGAPAKYSITFTGKWSQAFPKQY 60  
Db 1 MENPSPAALGKALCALLLATLGAAGQPLGSGESCSAGAPAKYSITFTGKWSQAFPKQY 60  
Qy 61 PLRFPAPQWSLLGAHSSDYSWKKQYVNSGLRDFAEERGEAWALKEITAEAGELQSV 120

Db 61 PLFRPPAQWSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
Qy 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Db 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Qy 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Db 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Qy 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Db 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Qy 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331  
Db 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331

## RESULT 2

US-10-616-279-2  
; Sequence 2, Application US/10616279  
; Publication No. US2004002307A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a No. US2004002307A1el RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSD1  
; CURRENT APPLICATION NUMBER: US/10/616,279  
; PRIOR FILING DATE: 2003-07-08  
; PRIOR APPLICATION NUMBER: US 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: US 09/732,357  
; PRIOR FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-616-279-2

Query Match 100.0%; Score 1760; DB 15; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1.6e-150;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MENPSPAAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
Db 1 MENPSPAAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
Qy 61 PLFRPPAQWSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
Db 61 PLFRPPAQWSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
Qy 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Db 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Qy 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Db 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Qy 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Db 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Qy 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331  
Db 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331

RESULT 3  
US-10-624-884-2  
; Sequence 2, Application US/10624884  
; Publication No. US20040152139A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: DNA Encoding a Novel RG-1 Polypeptide  
; FILE REFERENCE: 51791AUSC1  
; CURRENT APPLICATION NUMBER: US/10/624,884  
; PRIOR FILING DATE: 2003-07-22  
; PRIOR APPLICATION NUMBER: US 60/172,370  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: US 09/732,357  
; PRIOR FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-624-884-2  
Query Match 100.0%; Score 1760; DB 16; Length 331;  
Best Local Similarity 100.0%; Pred. No. 1.6e-150;  
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MENPSPAAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
Db 1 MENPSPAAALGKALCALLATLGAAGQPLGGESICSAGAPAKYSITFTGKWSQTAPPKQY 60  
Qy 61 PLFRPPAQWSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
Db 61 PLFRPPAQWSSLLGAHSSDYSMWRKNQYVNSGLRDFAEERGEAWALMKEIEAAGEALQSV 120  
Qy 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Db 121 HAVFSAFAPVSGTGTSAELEVRHSLVSFVRIVPSDFWVGVDSLDLDCDGRWREQA 180  
Qy 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Db 181 ALDLYPYDAGTDSGFTFSSPNFATIPQDVTVEITSSSPSHPANSFYYPRLKALPPIARVT 240  
Qy 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Db 241 LVRLRQSPRAFIIPAPVLPFRSDNEIVDSASVPETPLDCEVLSWSSWGLCGHCGRLGTSK 300  
Qy 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331  
Db 301 RTRVYRVQPNNGSPCPPELEEEAECPDNCV 331

## RESULT 4

US-10-895-183-2  
; Sequence 2, Application US/10895183  
; Publication No. US20050019845A1  
; GENERAL INFORMATION:  
; APPLICANT: Harkins, Richard  
; APPLICANT: Parkes, Deborah  
; APPLICANT: Parry, Gordon  
; APPLICANT: Schneider, Douglas  
; APPLICANT: Steinbrecher, Renate  
; TITLE OF INVENTION: RG1 Antibodies and Uses Thereof  
; FILE REFERENCE: 51791BUSM1  
; CURRENT APPLICATION NUMBER: US/10/895,183  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: US 60/489,032  
; PRIOR FILING DATE: 2003-07-22  
; NUMBER OF SEQ ID NOS: 31